MVA2000 IAPR Workshop on Machine Vision Applications, Nov. 28–30, 2000, The University of Tokyo, Japan

8—14

Active Contours for Bolus Tracking in X-Ray Images Sequences

Emmanuel Dellandrea, Pascal Makris, Nicole Vincent *
Laboratoire d'Informatique LI/E3I
Université François Rabelais de Tours

Michele Boiron†
Laboratoire de Physiologie et Motricité Digestive
UFR Médecine - Université François Rabelais de Tours

Abstract

This paper describes a tracking method that we have developed in order to track a *barium bolus* in radiological images sequences. Our method combines both region tracking by the means of a deformable region model, and contour tracking by the means of active contours. It allows the tracking of a *barium bolus* during its descent along the oesophagus and the lower oesophageal sphincter. Thus, we are enable to realize measurements of parameters that caracterize the sphincter, such as its length an its diameter.

This study takes part in a larger project, the aim of which is the improvement of the gastroesophageal reflux phenomenon understanding [2].

1 Introduction

Our goal is to track the *bolus* through radiological images sequences. The *bolus* can be considered as a deformable shape. Many tracking algorithms have been proposed lately [3, 4, 5]. The method we have developed combines region tracking thanks to a deformable region model, and contour tracking thanks to active contours. The tracking algorithm is presented figure 1.

The first step of the algorithm is the identification of the object to track: region R (the *bolus*). To do so, the first image of the sequence, where the *bolus* appears, has to be determined thanks to the study of grey levels variations. Then, the **active contours** algorithm is applied in order to obtain the boundary of the region R.

Once the region R has been detected, it then has to be tracked. The tracking method is composed of four main modules:

- **prediction** of the motion parameters thanks to the **Kalman filter**
- **optimization** of the predicted region thanks to a **deformable region model**
- **optimization** of the region boundary thanks to **active contours**
- **motion estimation** thanks to a **simulated annealing** method

These four steps will be described, after a presentation of the **motion modelization**.

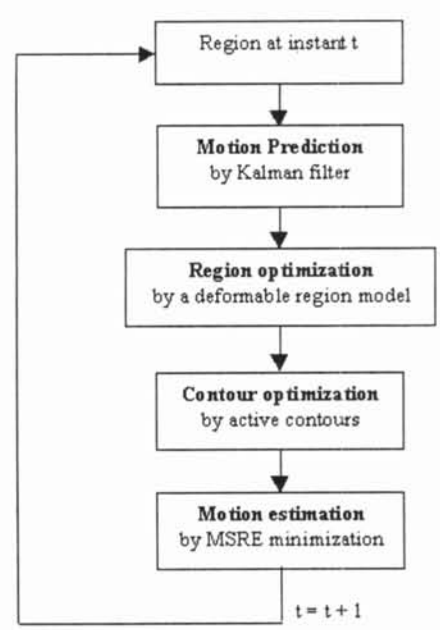

Figure 1: Tracking algorithm.

2 Motion modelization

In order to modelize the motion of the region R, we have used an affine motion model with six param-

*Address: 64 avenue Jean Portalis 37000 Tours France. E-mail: dellandrea@univ-tours.fr, makris@univ-tours.fr, vincent@univ-tours.fr

†Address: 2 bis boulevard Tonnellé 37000 Tours France. E-mail: boiron@med.univ-tours.fr

eters proposed in [9]. The displacement $\vec{d}$ of each point (x, y) of the region R is given by:

$$\vec{d}_R \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} t_x \\ t_y \end{pmatrix} + \begin{pmatrix} a & c \\ b & d \end{pmatrix} \begin{pmatrix} x - x_G \\ y - y_G \end{pmatrix}$$

The vector $\begin{pmatrix} t_x \\ t_y \end{pmatrix}$ represents the translation parameters of the gravity center, the coordinates of which are: (x_G, y_G).

Thus, the motion is entirely defined by a vector called the **motion descriptor**:

$$\Theta = [t_x, t_y, a, b, c, d]$$

3 Motion parameters prediction by the kalman filter

The **Kalman filter** introduced in [7], and presented in [10], represents a recursive solution to the discrete-data linear filtering problem. It consists in the estimation of a state vector Ψ of a system governed by the equation:

$$\Psi_{t+1} = A_t \Psi_t + \omega_t$$

from noisy measurement s so that:

$$s_t = H_t \Psi_t + v_t$$

The Matrix A represents the transition matrix from time t to $t+1$. H is called the observation matrix. ω and v are random variables representing respectively the process noise and measurement noise.

We have used the Kalman filter in order to predict an optimal **motion descriptor** thanks to previous measurements. Thus, our state vector is composed of the motion parameters and their first order temporal derivatives, the motion being considered as non-accelerated:

$$\Psi_t = (t_{x_t}, t_{y_t}, a_t, b_t, c_t, d_t, \dot{t}_{x_t}, \dot{t}_{y_t}, \dot{a}_t, \dot{b}_t, \dot{c}_t, \dot{d}_t)^T$$

Only motion parameters are measured. Therefore, the measurement vector s is:

$$s_t = (t_{x_t}, t_{y_t}, a_t, b_t, c_t, d_t)^T$$

The filter provides an estimation of the state vector Ψ_t, at instant t, using noisy measurements s_t, in order to predict the state at instant $t+1$: Ψ_{t+1}.

4 Optimization of the region by a deformable region model

Once the prediction phase has been realized, the predicted motion parameters are then available. A predicted contour is thus obtained. However, with this prediction, the region does not always perfectly match the real region. This is the reason why we have used a deformable region model, inspired from the one of [3]. The aim of this model is to allow the optimization of the region shape and position.

The deformable region model is defined by both a contour and an energy.

The optimization of the deformable region is realized iteratively by a minimization of its energy, measuring the similarity of the grey levels, between the points of the region and those of the region extracted from the previous image.

The use of this information given by the whole region improves the accuracy of the tracking in both case of large displacements of the object and noisy images.

The region deformation is entirely constrained by the motion model (affine model). Thus, the deformable region parameters that are optimized are the region motion parameters.

The **region contour** is modelized by an ordered set of n points:

$$\omega = \{\omega_i = (x_i, y_i)/i \in [1, n]\}$$

The points inside the region are obtained thanks to a filling algorithm, following a linear interpolation of the contour points.

The **region energy** is defined by:

$$E = \frac{1}{S_{R_t}} \sum_{\vec{p} \in R_t} [I_{t+1}(\vec{p} + \vec{d}_t^+(\vec{p})) - I_t(\vec{p})]^2$$

where:
R_t is the region extracted from the previous image at instant t,
S_{R_t} is the surface of the region R_t in pixels,
$\vec{p}$ is a point located in the image,
$d_t^+(\vec{p})$ is the translation vector of the point $\vec{p}$ from instant t to instant $t+1$,
$I_t(\vec{p})$ is the grey level of the image at instant t and at point $\vec{p}$.

This energy only depends on $\vec{d}_t^+(\vec{p})$ and so Θ_t^+.

Thus, minimising the deformable region energy amounts to determine the motion descriptor Θ_t^+ for which this energy is minimum:

$$\Theta_t^+ = \arg_{\Theta_t^+} \min E$$

The minimization of the energy is realized thanks to a **simulated annealing** method described in [6]. The advantage of this method is that, contrary to traditionnal methods such as the *steepest descent* [9], it does not give the first minimum encountered, but, generaly the global minimum.

5 Optimization of the boundaries by active contours

The previous step of the algorithm has allowed the optimization of the region that was predicted by the **Kalman filter**. However, deformable region models used for region tracking, have difficulty in determining the region boundary with accuracy. Thus, it has appeared interesting to associate a deformable region model with a deformable contour model. The last-mentioned goal is to make the object boundary match the real object boundary. The deformable contour model chosen is the **active contours**.

Several methods of actives contours could be conceivable [8, 1]. The one proposed in [11] has been chosen since it has appeared to be stable and fast. This method relies on the representation of the contour by an ordered set of n points, which, in our case, is the same as the one of the deformable region model:

$$\omega = \{\omega_i = (x_i, y_i)/i \in [1, n]\}$$

These points will approach iteratively the boundary of the object by a minimization of an energy function.

For each point ω_i' in the neighborhood of ω_i, an energy term is computed:

$$E(\omega_i') = \alpha E_{int}(\omega_i') + \beta E_{ext}(\omega_i')$$

where E_{int} is the internal energy term that imposes a certain regularity of the contour and E_{ext} is the external energy term linked to the global image. α and β are constants providing the relative weight of the energy terms.

Each point ω_i is then moved to its neighbor corresponding to the location of the minimum value of E.

The choice of the energy terms depends on the object to segment and on the image type.

Three different terms in the internal energy have been introduced: a continuity one, that encourages the equidistance between the contour points, a curvature one, that encourages smoothness of the contour, and a balloon one, that drives the contour to swell up to objects boundaries. The last-mentioned is usefull, in our images, when the ***barium bolus*** expands.

Moreover, in the external energy two terms have been introduced: a magnitude term, attracting the contour to edges in the image, and an intensity term, attracting the contour to low intensity regions (the *bolus* is a dark area in our images).

Once the contour ω has been optimized, it then has to be re-sampled. The aim of this step is to improve the accuracy of the contour by imposing the contour points to be separated by the same distance. Besides, the number of the contour points can evolve in order to fit the dimension of the deformable object.

This step plays an important part since the deformable region model and the active contours models relies on the representation of the contour of the region by the same set of points ω. Thus, it is necessary for the points of ω to be homogeneously distributed around the region contour.

In order to re-sample the contour, a linear interpolation between the points of ω is realized. This enables to determine the set of pixels in the image that belong to the region contour. They represent a one-pixel-wide curve. The new contour ω' is extracted from that curve so that the points of ω' are separated by the same number of pixels.

6 Motion estimation

At this point, the contour ω_{t+1} of the region R_{t+1} in the image I_{t+1}, is clearly defined.

Finally, it is necessary to estimate the motion from I_t to I_{t+1} in order to supply the **Kalman filter**.

The principle of the **motion estimation** step is to minimize the $MSRE$ (*Mean Squared Reconstruction Error*) of the region R_{t+1} in I_{t+1} from the region R_t in I_t:

$$MSRE_{R_{t+1}} = \frac{1}{S_{R_{t+1}}} \sum_{\vec{p} \in R_{t+1}} [I_t(\vec{p}+\vec{d}_{t+1}^{-}(\vec{p})) - I_{t+1}(\vec{p})]^2$$

where:
$d_{t+1}^{-}(\vec{p})$ is the translation vector of the point $\vec{p}$ from instant $t+1$ to instant t.

This error plays the same part than the deformable region model energy, but the temporal axis is considered in the opposite direction (from t to $t+1$ for the region model, and from $t+1$ to t for the motion estimation).

The $MSRE$ only depends on $\vec{d}_{t+1}^{-}(\vec{p})$ and so Θ_{t+1}^{-}.

Thus, minimising the deformable region energy amounts to determine the motion descriptor Θ_t^{-} for which the energy is minimum:

$$\Theta_{t+1}^{-} = \arg_{\Theta_{t+1}^{-}} \min MSRE_{R_{t+1}}$$

The minimization is also realised by a **simulated annealing** method.

7 Experimental Results

We dispose, for the time being, of four sequences acquired on two voluntaries. The results of the *bolus* tracking on one sequence are presented in figure 2.

The bright contour corresponds to the contour given after the region optimization by the deformable region model. The dark contour corresponds to the optimized and re-sampled contour given by active contours.

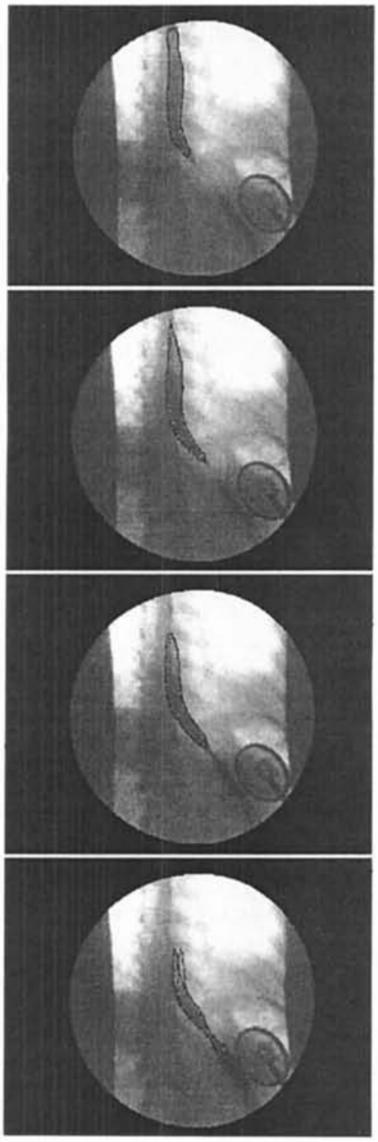

Figure 2: Results of the bolus tracking at instants $t = 46$, $t = 59$, $t = 66$ and $t = 73$ of a sequence.

8 Conclusion

A tracking method, that combines both region and contour tracking, has been presented in this paper. Its aim is to realize a *barium bolus* tracking in radiological images sequences.

The results have shown the efficiency of the method, especially when images contrast is sufficient.

The tracking allows the measurements of parameters that caracterize the lower oesophageal sphincter, such as its length and its diameter. The fusion of these results with the ones of a swallowing sounds analysis will allow a better understanding of both the gastroesophageal reflux phenomenon and the origins of the sounds generated during barium swallowing [2].

At present, the means of investigation are radiological and manometric explorations. Thanks to our work, a non-invasive and atraumatic method entirely based on sounds recording, could be conceivable.

References

[1] A. A. AMINI, T. E. WEYMOUTH and R. C. JAIN, *Using dynamic programming for solving variational problems in vision*, IEEE Transactions on pattern analysis and machine intelligence, Vol. 12, No. 9, pp. 855-867, 1990.

[2] M. BOIRON, B. ATIPO, L. PICON, P. ROULEAU and EH. METMAN, *Esophageal swallowing phase assessed by audiosignal recording: Relationship with manometry in gastroesopheal reflux disease*, Dig. Dis. Scix., 44, pp. 529-545, 1999.

[3] B. BASCLE and R. DERICHE, *Region tracking through image sequences*, Rapport de recherche 2439, INRIA-Sophia Antipolis, 21 pages, Décembre 1994.

[4] L. BONNAUD and C. LABIT, *Etude d'algorithmes de suivi temporel de segmentation basée mouvement pour la compression de séquences d'images*, Publication interne 793, IRISA-Université de Rennes, 43 pages, Janvier 1994.

[5] L. BONNAUD, *Schémas de suivi d'objets vidéo dans une séquence animée: application à l'interpolation d'images intermédiaires*, Thèse de doctorat, IRISA-Université de Rennes, 1998.

[6] A. CORANA, M. MARCHESI, C. MARTINI and S. RIDELLA, *Minimizing multimodal variables with the "Simulated annealing" algorithm*, ACM Transactions on Mathematical Software, Vol. 13, No. 3, pp. 262-280, September 1987.

[7] R. E. KALMAN, *A new approach to linear filtering and prediction problems*, Transactions of the ASME- Journal of Basic Engineering, pp. 35-45, March 1960.

[8] M. KASS, A. WITKIN and D. TERZOPOULOS, *Snakes : Active contour models*, International Journal of Computer Vision, Vol. 1, pp. 321-331, 1998.

[9] H. NICOLAS, *Hiérarchie de modèles de mouvement et méthodes d'estimation associées. Application au codage de séquences d'images*, Thèse de doctorat, IRISA-Université de Rennes 1, 1992.

[10] G. WELCH and G. BISHOP, *An introduction to the Kalman Filter*, *http://www.cs.unc.edu/ ~welch/kalman/kalman_filter/kalman.html*, 1997.

[11] D.J. WILLIAMS and M. SHAH, *A fast algorithm for active contours and curvature estimation*, CVGIP: Image Understanding, Vol. 55, No. 1, pp.14-26, January 1992.